Towards an understanding of dimensions, predictors, and gender gap in written composition

Young-Suk Kim[1], **Stephanie Al Otaiba**[2], **Jeanne Wanzek**[1], and **Brandy Gatlin**[3]

[1]Florida State University and Florida Center for Reading Research

[2]Southern Methodist University

[3]Florida State University

Abstract

We had three aims in the present study: (1) to examine the dimensionality of various evaluative approaches to scoring writing samples (e.g., quality, productivity, and curriculum based writing [CBM]) , (2) to investigate unique language and cognitive predictors of the identified dimensions, and (3) to examine gender gap in the identified dimensions of writing. These questions were addressed using data from second and third grade students ($N = 494$). Data were analyzed using confirmatory factor analysis and multilevel modeling. Results showed that writing quality, productivity, and CBM scoring were dissociable constructs, but that writing quality and CBM scoring were highly related ($r = .82$). Language and cognitive predictors differed among the writing outcomes. Boys had lower writing scores than girls even after accounting for language, reading, attention, spelling, handwriting automaticity, and rapid automatized naming. Results are discussed in light of writing evaluation and a developmental model of writing.

Keywords

Dimensionality; Writing Quality; Writing Productivity; CBM; Gender

Students' writing skill is assessed in multiple ways. To assess a discourse level writing skill (e.g., ability to writing in paragraphs), students are typically asked to write written compositions, and written compositions are evaluated using multiple approaches such as writing quality, writing productivity, or curriculum-based measurement (CBM) writing scoring. Another widely used writing assessment measures a sentence level writing ability by asking students to produce grammatically correct sentences within a specified time (Writing Fluency task of the Woodcock Johnson Tests of Achievement-III [WJ-III], Woodcock, McGrew, & Mather, 2001). Despite the existence of various ways of assessing students' writing skill, researchers and practitioners have a limited understanding of how these various assessments and evaluative approaches are related and whether they tap into or capture similar or dissociable dimensions of writing. A clearer understanding of assessment approaches is needed to advance theories of development and to guide practitioners in using

Corresponding Author Young-Suk Kim Associate Professor Florida State University & Florida Center for Reading Research Florida State University 1114 W. Call St. Tallahassee, FL 32306 ykim@fcrr.org.

assessment data to inform instruction and intervention. In the present study we address this question with three goals. First, we examined how various approaches to writing assessments converge or diverge into different dimensions, using various evaluative approaches such as writing quality, productivity, and CBM scoring as well as using a widely used sentence level task, the WJ-III Writing Fluency task. Second, we further examined how language and cognitive skills relate to the identified dimensions. Finally, given the consistent achievement gaps between boys and girls on national writing assessments (e.g., National Center for Education Statistics, 2003), we also sought to examine gender differences across the identified dimensions of writing.

Approaches to Writing Evaluation

According to the simple view of writing (Juel, Griffith, & Gough, 1986), two necessary components of writing are ideation (i.e., generation and organization of ideas) and transcription skills. The first component, ideation, refers to the quality of ideas represented in writing, which is an essential, and arguably the most important, aspect to be evaluated in written compositions. Not surprisingly, writing quality has long and widely been examined in previous studies. Two key indicators of writing quality appear to be the extent of development and organization of ideas (Bereiter & Scardamalia, 1987; Juel et al., 1986). In fact, idea development and organization have been widely examined as indicators of writing quality in previous studies (Graham, Harris, & Chorzempa, 2002; Graham, Harris, & Mason, 2005; Kim et al., 2011; Kim et al., 2013; Olinghouse, 2008). Other widely used assessments of writing examine similar aspects. For example, the Test of Written Language-4th Edition (TOWL-4) includes theme development and organization, and another widely used writing evaluation approach in U.S. schools (Gansle et al., 2006), the 6+1 Trait rubric, includes idea development and organization/structure aspects in addition to other aspects such as word choice, sentence fluency, voice, presentation, and conventions.

The other component of the simple view of writing, transcription skill, allows generated ideas to be produced in written text and facilitates idea generation and development (see below; Berninger et al., 1997; Graham et al., 1997; Graham, Harris, & Fink, 2000; Kim et al., 2011). Therefore, the amount of written composition is constrained by transcription skills to a large extent, particularly for beginning writers. Not surprisingly, writing productivity is another widely examined dimension of writing (e.g., Abbott & Berninger, 1993; Berman & Verhoevan, 2002; Kim et al., 2011; Mackie & Dockrell, 2004; Olinghouse & Graham, 200; Scott & Windsor, 2000). Note that although the term, writing fluency, has been used often to refer to a similar construct, we use the term, writing productivity, because we are specifically referring to the amount of text produced, not the automaticity, effortlessness, and coordination of multiple processes which are defining characteristics of fluency (Berninger et al., 2010; LaBerge & Samuels, 1974). In addition, writing fluency has been conceptualized to refer to CBM writing (Ritchey et al., in press). Although the amount of text alone is not generally considered a yardstick or goal of good writing, good written composition requires a certain amount of text for the ideas to be sufficiently developed and articulated. Previous studies have examined writing productivity, and it has been shown to be a dissociable dimension from writing quality (Kim et al., 2014; Wagner et al., 2011), and correlations between writing quality and productivity tend to be fairly strong for children in

the elementary grades (e.g., .65 ≤ *rs* ≤ .82; Abbott & Berninger, 1993; Kim et al., 2014; Olinghouse & Graham, 2009). Writing productivity is measured using various indicators such as the total number of words, number of ideas, number of different words, and/or number of sentences (Kim et al., 2014; Kim, Park, & Park, 2013; Puranik, Lombardino, & Altmann, 2008; Wagner et al., 2011).

A third evaluative approach to writing employed in the present study is CBM scoring. CBM writing scoring includes some unique evaluative tools not included in the writing quality and productivity indicators noted above. Along with reading and math CBM measures, CBM writing measures are considered global outcome measures, or indicators, of students' overall writing performance (Deno, 1985) that are intended to signal whether the student needs further diagnosis and intervention. CBM writing measures were initially developed to screen and monitor progress in writing skills for students at risk for writing difficulty. Students are typically asked to write for 3 to 5 minutes in response to prompts (Coker & Ritchey, 2010; McMaster, Du, & Pestursdottir, 2009; McMaster et al., 2011), and their writing is evaluated using various scoring tools such as number of words written, correct word sequences (two adjacent words that are grammatically correct and spelled correctly), incorrect word sequences, words spelled correctly, percent of correct word sequences, and correct minus incorrect word sequences (see Graham et al., 2011; McMaster & Espin, 2007 for a review). Note that number of words written is not unique to the CBM writing scoring as it has been used as an indicator of writing productivity.

CBM writing measures have been shown to be reliable, and students' scores on CBM writing tend to be related to other writing measures with validity coefficients in the moderate range (see Graham et al., 2011 and McMaster & Espin, 2007 for a review; McMaster et al., 2009; Lembke, Deno, & Hall, 2003). In particular, the correct minus incorrect word sequences (CIWS) score tends to be the most strongly related to other writing measures with coefficients ranging from .60 to .75 (Espin et al., 2000; Espin, Weissenburger, & Benson, 2004). Recently, the percent of correct word sequences (%CWS), along with the CIWS, has also been shown to be highly ($r = .61$) related to a normed writing task (Test of Written Language-3) for children in middle school (Amato & Watkins, 2011).

Despite the reliability and validity evidence for CBM writing scoring procedures described in these previous studies, it is not clear how CBM writing scores should be conceptualized in terms of dimensionality. That is, do CBM writing scores capture dimensions such as writing quality or writing productivity, or do they measure a separate, overall global outcome measure of writing? Recently, CBM writing measures have been described as 'writing fluency', which is defined as the ease with which an individual "*produces* written text", and includes both "*text generation* (translating ideas into words, sentences, paragraphs, and so on) and *transcription* (translating words, sentences, and higher levels of discourse into print)." (emphasis in the original text, Ritchey et al., in press). A critical question is whether potential writing fluency indicators capture a dissociable dimension, apart from other widely examined dimensions such as writing quality and productivity. Although its theoretical foundations is still in its nascent stage, we included CBM writing scores in the present study because of validity evidence with other writing measures, and its

potential practical utility for progress monitoring purposes as CBM indicators have been shown to be sensitive to growth over time within a short time period (e.g., 2 weeks; see Espin et al., 2004; McMaster & Espin, 2007).

Finally, although writing skill is typically assessed by asking the child to produce a written composition, other tasks also have been used. One such widely used standardized subtest is the Writing Fluency task of the WJ-III (Woodcock et al., 2001). This task assesses sentence-level, rather than paragraph-level writing. Children are presented with a picture and three words, and they are asked to write a sentence about the picture using the three words. The child's score is the number of correct and meaningful written sentences based on the three words that were presented. However, how the WJ-III Writing Fluency relates to other dimensions of writing is an open question.

In the present study, we examined dimensionality of writing using children's data from written compositions as well as the Writing Fluency task of the WJ-III. Children's written compositions were evaluated by indicators of writing quality, productivity, and CBM writing scores. For the Writing Fluency task of the WJ-III, scores following the WJ-III scoring guidelines were used. Our goal in the present study was to extend our understanding of writing dimensionality. Previous studies have shown that writing quality, productivity, spelling and writing conventions, and syntactic complexity are dissociable dimensions for typically developing children in grades 1 and 4, and children with language impairments (Kim et al., 2014; Puranik, Lombardino, & Altmann, 2008; Wagner et al., 2011). In the present study, we expand this line of research by examining how CBM scores and the Writing Fluency task of the WJ-III are related to writing quality and productivity dimensions using data from children in grades 2 and 3.

Predictors of Writing Skills

As noted above, writing is composed of at least two components skills: transcription skills and ideation (Berninger & Swanson, 1994; Juel et al., 1986). Transcription skills such as spelling and handwriting allow mental resources such as attention and working memory to be available for idea generation and translation processes (Berninger & Swanson, 1994; Graham, 1990; Graham, Harris, & Fink, 2000; Graham et al., 1997; Scardamalia, Bereiter, & Goleman, 1982). Much evidence supports the role of transcription skills in writing (Berninger, 1999; Graham et al., 1997; Jones & Christensen, 1999; Kim et al., 2011, 2014; Wagner et al., 2011). Handwriting skill is typically assessed by asking the child to write alphabet letters or to copy sentences or paragraphs as accurately and fast as possible within a specified time (e.g., Abbott & Berninger, 1993; Graham et al., 1997; Kim et al., 2011; Wagner et al., 2011)

Although ideation, the other component of writing according to the simple view of writing, is challenging to directly measure, it has been largely measured by means of oral language use (e.g., Chenoweth & Hayes, 2003; Hayes, 2012). Generated ideas cannot be produced without being translated into oral language because the child has to express ideas using appropriate words, encode them using appropriate syntactic structure, and organize and present them in a logical sequence. Therefore, oral language proficiency would determine

how the generated ideas are adequately expressed. Evidence of the importance of oral language in written composition is accumulating from beginning writers to those in middle school (Berninger & Abbott, 2010; Kent et al., 2014; Kim et al., 2011, 2013, 2014; Olinghouse, 2008) as well as children with language impairment (Dockrell, Lindsay, & Connelly, 2009; Dockrell & Connelly, in press; Kim, Puranik, & Al Otaiba, 2014; Puranik, Lombardino, & Altmann, 2007). Given that writing is a production or constructed-response task, children's transcription skills constrain the extent to which generated ideas can be transcribed into generated text (Berninger et al., 2002; Juel et al., 1986).

In addition to the above noted skills, the not-so-simple view of writing states that executive function processes such as attention, planning, self-regulation, and working memory are critical supports for writing development (Berninger & Winn, 2006). Attention, in particular, has been shown to be related to writing for children in first and second grade (Hooper et al., 2002, 2011; Kent et al., 2014; Kim et al., 2014). Additional evidence underscoring the importance of attention in writing comes from studies with children who have Attention Deficits or Attention Deficit Hyperactivity Disorder (ADHD); converging evidence suggests that students with ADHD made more spelling and grammatical errors (Casas, Ferrer, & Fortea, 2013; Gregg, Coleman, Stennett, & Davis, 2002; Re, Pedron, & Cornoldi, 2007), made more content errors or digressions, and demonstrated weaker text structure features than children without ADHD (Casas et al., 2013).

Individual differences in reading also have been shown to matter for children's writing development (Shanahan, 2006). Studies have shown that reading comprehension was related to written composition quality and productivity for children in elementary and middle school grades (Berninger & Abbott, 2010; Berninger et al., 2002; Kim et al., 2013, 2014). Children's reading ability might influence written composition skill via reading experiences. Greater reading ability and consequent text reading might allow the opportunity for the child to acquire vocabulary and syntactic structures, and organization of written text as well as content (Berninger et al., 2006). In fact, children with impaired reading comprehension had weaker story content and organization in their writing (Cragg & Nation, 2006).

Writing involves juggling of multiple processes to even greater extent than in reading. Therefore, the ability to coordinate multiple aspects is likely to be important. Some previous studies have examined Rapid Automatized Naming (RAN) in this regard as a potential predictor of writing. Numerous studies have shown that rapid automatized naming is related to reading (Compton, DeFries, & Olson, 2001; de Jong, & van der Leij, 2003; Kim, 2011; Kirby, Parrila, & Pfeiffer, 2003; Savage et al., 2005; Wolf & Bowers, 1999; Wolf &O'Brien, 2001). However, despite a robust relation to reading in various languages, researchers differ about what it exactly measures, and hypotheses include phonological processing (Wagner & Torgesen, 1987), automaticity of processes (Bowers, 1995; LaBerge & Samuels, 1974; Spring & Davis, 1988), global processing speed (Kail & Hall, 1994), and multiple constructs such as lexical access, automaticity, attentional, visual, and articulatory processes (Wolf & Bowers, 1999). If RAN measures automaticity of processes, its influence might largely overlap with that of automaticity of transcription skills, and thus may not be related to writing over and above transcription skills. In contrast, if RAN captures multiple constructs beyond what is captured by transcription skills, it would be related to writing over

and above transcription skills. Although RAN has not been examined for young English speaking children, there is some emerging evidence from studies with Chinese children that suggests that RAN is related to writing (Chan, Ho, Tsang, Lee, & Chung, 2006; Ding, Richman, Yang, & Guo 2010; but see Yan et al., 2012).

Gender and Writing

Gender appears to matter in children's writing achievement. Girls have outperformed boys in writing consistently across grades ever since writing was included in the National Assessment of Educational Progress (NAEP). For instance, in 2002 in which writing was assessed including children in grade 4 as well as those in grades 8 and 12, girls outperformed boys in all the three grades with gaps ranging from 17 to 25 points, (National Center for Education Statistics, 2003). Similarly, gender gaps have been reported for children in elementary grades (Berninger & Fuller, 1992; Knudson, 1995). Despite these consistent gender gaps in writing, our understanding about gender gaps in writing and potential sources of gender gaps is limited, particularly for children in elementary grades. One potential source of gender gaps seen in older students is their attitude toward writing. Among adolescents, males tend to have less positive attitudes toward writing than do females (Knudson, 1992; Pajares & Valiante, 1999), and see less value in writing and express less satisfaction with writing activities (Lee, 2013). Studies of younger students reported mixed findings about the relation of attitude toward writing and children's writing skill. Knudson (1995) investigated gender and writing attitude with children in grades 2 and 6 and showed that children's attitude toward writing predicted their writing skill. In contrast, a study with children in grades 1 and 3 revealed that girls had more positive attitudes than boys toward writing as early as in grade 1, but this difference was not related to their writing skill (Graham, Berninger, & Fan, 2007).

Another potential source of gender gaps in writing achievement is reading or reading-related skills. As noted earlier, evidence suggests that reading is one of the component skills of writing. Evidence also indicate that male students have been consistently outperformed by female students in reading (e.g., National Center for Education Statistics, 2011), and a greater number of boys are identified with reading disabilities (Hawke et al., 2009; Miles, Haslum, & Wheeler, 1990; Yoshmasu et al., 2013; but see Shaywitz, Shaywitz, Fletcher, & Escobar, 1990). Therefore, differences in reading or reading-related skills might explain differences in writing skills between boys and girls. Furthermore, boys in grades 1, 2, and 3 had lower scores in another writing component skill, transcription skill (Berninger & Fuller, 1992). In the present study, we examined whether gender differences were found for children in grades 2 and 3 in the identified writing dimensions, and if so, to what extent gender differences were explained by the included language and cognitive skills (e.g., reading, attention, and transcription skills).

Present Study

The primary goal of the present study was to examine the dimensionality of various writing evaluation approaches, predictors of various dimensions, and the gender gap in writing. Specific research questions were as follows.

1. What are the relations of CBM writing measures (i.e., CIWS and %CWS) and the WJ-III Writing Fluency task to writing quality and writing productivity indicators? Are CBM writing measures and the WJ-III Writing Fluency task measure dissociable dimensions from writing quality and writing productivity?

2. How are language and cognitive skills related to the identified writing dimensions?

3. Are there performance differences between boys and girls in the identified writing dimensions (e.g., writing quality and productivity) after accounting for children's language and cognitive skills?

To address these questions, we used data from second and third grade children ($N = 494$) who were administered multiple writing tasks: written compositions in response to three prompts (one normed task, and two experimental tasks), and a sentence-level task, the WJ-III Writing Fluency task. Students' compositions were evaluated using a variety of approaches including writing quality indicators such as idea development and organization, writing productivity indicators such as number of words written and number of ideas, CBM writing scoring such as CIWS and % CWS, and scoring protocols in the standardized tasks. Language and cognitive skills included oral language, reading, transcription (spelling and handwriting fluency), attention, and rapid automatized naming.

We hypothesized that writing quality and productivity would be dissociable dimensions based on previous studies (Kim et al., 2014; Puranik et al., 2008; Wagner et al., 2011). We also hypothesized that the CBM writing scores would be a dissociable construct because although validity coefficients of CIWS and %CWS were acceptable, they are not extremely highly correlated with other writing measures (e.g., Amato & Watkins, 2011; McMaster & Espin, 2007). In contrast, we did not have a priori prediction about the WJ-III Writing Fluency task. It was also hypothesized that various language and cognitive skills would be differentially related to different writing outcomes based on a prior study (Kim et al., 2014). Finally, gender differences were hypothesized, and language and literacy skills were expected to explain gender differences to some extent.

Method

Participants and Sites

Students in the present study included 494 children in grades 2 (mean age = 7.80) and 3 (mean age = 8.82). These students were drawn from 76 classrooms in 10 schools in a mid-sized city. The students were 51.2% male and 76.1% received free and reduced price lunch. Six of the 10 schools were Title I schools, indicating that the majority of the students in the school were eligible for the Free or Reduced Price Lunch program. Students' racial backgrounds were as follows: 60% African Americans, 29% Whites, and the rest were composed of Asians and multiracial children. The students and their families had consented for their participation and all guidelines for human research protection continued to be followed in the present study.

Measures

Writing tasks—Four tasks were used to assess children's written composition skill: two standardized and normed tasks, and two experimental prompts. The first task was the Writing Fluency subtest of the Woodcock-Johnson Tests of Achievement – 3rd edition (WJ-III; Woodcock et al., 2001). In this subtest, students were provided with a series of pictures and three corresponding words and were instructed to write a sentence about the picture that included the words given. Students were given seven minutes to complete as many sentences as they could. For the scoring of this subtest, we used standard scoring procedures outlined in the testing manual. Namely, students received one point for each complete sentence. In order to receive credit, the sentence had to be clear in meaning and include critical words to make the sentence reasonable. Students were not penalized for errors in punctuation, spelling, capitalization, or for poor handwriting. Using the Rasch analysis procedure, the reliability coefficient was reported to be .72 for 7- and 8-year-olds (McGrew, Schrank, & Woodcock, 2007).

We also asked children to write on three prompts: one prompt from the Written Essay test of the Wechsler Individual Achievement Test-Third Edition (WIAT-III; Wechsler, 2009) and two experimental prompts, one narrative prompt and one expository prompt. The WIAT-III was selected as a widely-used writing assessment that could be compared to other research (e.g., see Berninger & Abbott, 2010). In the WIAT-III Essay task, children were asked to write about a favorite game and include at least three reasons as support. Note that standard scores in this task are available starting with children in grade 3, and not available for children in grade 2. Despite lack of standard scores, this task was deemed useful for children in grade 2 for the purpose of examining dimensionality and predictive relations. In addition, assessors confirmed that this topic did not appear to be difficult for children in grade 2.

The experimental narrative prompt was "One day when I got home from school…" Children were asked to write about any interesting events that occurred responding to the prompt (Kim et al., 2013, 2014; McMaster et al., 2009; McMaster et al., 2011). The experimental expository prompt was adapted from a previous study (Wagner et al., 2011). In this task, children were asked to write about a classroom pet they would like and explain why. For each prompt, children were given a 10-minute time limit.

Writing Evaluation

Children's written compositions for the WIAT Essay Composition task and two experimental prompts were evaluated on writing quality, writing productivity, and CBM writing measure scoring (see below). In addition, the WIAT essay was scored according to the examiner's manual (see below). Children's responses to the WJ-III Writing Fluency task were evaluated only according to the examiner's manual noted above because the responses were sentences, not passage level composition.

Writing quality scoring—The quality of children's written composition was evaluated on the extent to which their ideas were developed and the extent to which the ideas were presented in an organized manner, using a rating scale of 1 to 7. In this idea development aspect, high scores were given to compositions with rich and detailed ideas and ideas with

unique and interesting perspectives. In the organization aspect, compositions with logical sequence and transitions of expressed ideas with overall structure of beginning, middle, and end received high scores. These were similar to the 6-point scale version of the 6+1 Trait rubric, but were adapted to a 1 - 7 rating scale, representing low quality and high quality, respectively. When using 45 writing samples per prompt, reliabilities (Cohen's kappa) ranged from .82 to .88 for ideas and organization.

Writing productivity scoring—Two indicators were used for writing productivity: total number of words written and number of ideas. The number of words has been widely used as an indicator of compositional productivity in writing (e.g., Abbott & Berninger, 1993; Berman & Verhoevan, 2002; Kim et al., 2011; Mackie & Dockrell, 2004; Puranik et al., 2008; Scott & Windsor, 2000; Wagner et al., 2011). Words were defined as real words recognizable in the context of the child's writing despite some spelling errors. Random strings of letters or sequences were not counted as words. Random strings of letters were identified by comparing a record of what the child said she had written to her written composition. These were highly rare in the sample (less than 10). The number of ideas was a total number of propositions, which were defined as predicate and argument. For example, "I went upstairs and took a bath" was counted as two ideas (see e.g., Kim et al., 2011, 2013; Puranik et al., 2008). Repeated ideas were only counted once. When using 45 writing samples per prompt, reliabilities were .88 for the number of ideas (kappa) and .99 for the number of words (similarity).

Curriculum-based measure scoring—Each essay was individually analyzed for curriculum-based measures (CBM) including the Correct Word Sequence ("any two adjacent, correctly spelled words that are acceptable within the context of the sample," McMaster & Espin, 2007, p. 76), and the Incorrect Word Sequence ("any two adjacent letters that are incorrect," McMaster & Espin, 2007, p. 76). From these, a Correct minus Incorrect Word Sequence (CIWS) was obtained by subtracting incorrect words from correct word sequence. The percentage of correct word sequences (%CWS) was calculated by dividing the number of CWS by the total number of words written. In the data analysis, we used CIWS and %CWS for two reasons: (1) number of words written has been used as an indicator writing productivity and thus, not unique to CBM writing, and correct word sequence is highly related to the number of words written (because children who write more tend to have greater number of correct sequence); and (2) evidence indicates that CIWS and %CWS have greater validity coefficients with other writing tasks than the other CBM writing scoring (e.g., McMaster & Espin, 2007). Reliability for each type of scoring was established using 45 pieces per prompt. We used an equation that produced quotients to indicate the proximity of the coder's score for each measure to that of the primary coder (i.e., similarity coefficients; Shrout & Fleiss, 1979) and reliability for each measure ranged from .92 to .99.

WIAT standardized scoring—In addition to the above noted evaluative measures, students' compositions for the WIAT Essay Composition task were scored according to the manual. The WIAT scoring includes the total number of words, thematic development and text organization (theme and organization hereafter), and a supplementary score called

'grammatical score.' The grammatical score is highly similar to CIWS in CBM writing although slight differences are found in operationalization (e.g., WIAT does not give credit for titles or endings such as 'The End' whereas conventional CBM writing does). The unique scoring in the WIAT task, thus, is the theme and organization, and students' compositions were assigned scores in the following categories: Introduction, Conclusion, Paragraphs, Transitions, Reasons Why, and Elaborations. The maximum score possible for the theme and organization component was 20 points. Inter-rater reliability was established by having two independent coders score 50 essays and comparing individual points assigned. The number of agreements was divided by the total number of agreements plus disagreements, resulting in a reliability coefficient of .85. A standard score for theme and organization was computed for each student based on his or her chronological age at the time of testing. The standard score for the WIAT Essay Composition task is a composite of the standard score for theme and organization and for total number of words written.

Predictors

Predictors were selected based on our review of the literature and included oral language, reading, spelling, handwriting fluency (letter writing and story copying tasks), attention, and rapid automatized naming.

Oral language—Children's oral language skill was measured by the following three tasks: WJ-III Picture Vocabulary (Woodcock et al., 2001), the Test of Narrative Language Narrative Comprehension subtest (Gillam & Pearson, 2004), and the Oral and Written Language Scales Listening Comprehension subtest (Carrow-Woolfolk, 2011). In the Picture Vocabulary task, children were asked to identify pictured objects. Test-retest reliability is reported as .71 - .73 for 7- and 8-year-olds (McGrew et al., 2007). The Narrative Comprehension subtest of the Test of Narrative Language includes three individual tasks in which each student listens to a short story and is then asked to answer specific comprehension questions. The internal consistency of this subtest is .87 and test-retest reliability is .85 (Gillam & Pearson, 2004). In the Listening Comprehension subtest of the Oral and Written Language Scales, students listen to a stimulus sentence and are asked to point to one of four pictures that corresponds to the sentence read aloud by the tester. This subtest's reported split-half internal reliability ranges from .96 to .97 for the age group of our sample (Carrow-Woolfolk, 2011).

Reading—Children's reading skill was assessed using five measures: the WJ-III Letter Word Identification and Passage Comprehension subtests (Woodcock et al., 2001), the Sight Word Efficiency subtest of the Test of Word Reading Efficiency – Second Edition (Torgesen, Wagner, & Rashotte, 2012), the Oral Reading Fluency subtest of the WIAT-III (Wechsler, 2009), and the Test of Silent Reading Efficiency and Comprehension (TOSREC; Wagner, Torgesen, Rashotte, & Pearson, 2010). For the Letter Word Identification task, the child is asked to read aloud letters and words of increasing difficulty. For the WJ-III Passage Comprehension subtest, students are asked to silently read a short passage and provide a missing word that makes sense within the context of the passage. Reliabilities (test-retest) are reported to be .96 for both the Letter Word Identification and the Passage Comprehension subtests for students in the age range of the students we assessed (McGrew

et al., 2007). In the Sight Word Efficiency task, the child is asked to read words of increasing difficulty with accuracy and speed. Test-retest reliability for the Sight Word Efficiency is reported to be .93 for 6- and 7-year-olds and .92 for 8 – 12-year-olds. For the WIAT Oral Reading Fluency task, the child is asked to read two grade-level passages aloud. The student is timed during both readings and the completion time is recorded in seconds for each prompt. Each raw score is then used to compute an average weighted raw score to determine oral reading fluency. Test-retest reliability for the WIAT Oral Reading Fluency subtest is reported as .93. For the Test of Silent Reading Efficiency and Comprehension, the student is given three minutes to read a series of statements and determine if each statement is true or not. The authors report alternate-form reliability coefficients ranging from .87 - .95 for students in grades 2 and 3.

Spelling—Children's spelling skill was measured by a dictation task, the WJ-III Spelling subtest (Woodcock et al., 2001). Once a student misspells six consecutive words, the test is discontinued. The authors of this assessment report test-retest reliability coefficients of .91 and .88 for 7- and 8-year-olds respectively.

Letter writing automaticity—The WIAT-3 Alphabet Writing Fluency task was used, in which children were asked to write as many letters of the alphabet as possible with accuracy. This task assesses how well children access, retrieve, and write letter forms automatically. Research assistants asked children to write as many letters of the alphabet as they could in a 30-second time period. Children received a score for the number of correctly written letters. One point was awarded for each correctly formed letter. Interrater reliability (Cohen's kappa) for this subtest was .88 for our sample.

Story copying—Another transcription skill, the ability to copy letters, was measured by an experimental story copying task. In this task, students were instructed to copy a narrative story titled "Can Buster sleep inside tonight?" as fast as they could. The story had 519 words and involves a dog named Buster being muddy and being bathed so that he could sleep inside. Students were given one minute to write as much of the story verbatim as possible. Children received a score for the number of letters correctly formed, which was calculated as the difference between the number of letters attempted and the number of letter errors made. Interrater reliability (Cohen's kappa) for this measure was established at .91.

Attention—The first nine items of the Strengths and Weaknesses of ADHD-symptoms and Normal behavior scale (e.g., SWAN; Swanson et al., 2006) were used to measure children's attentiveness. SWAN is a behavioral checklist that includes 30 items that are rated on a seven-point scale ranging from a score of one (far below average) to seven (far above average) to allow for ratings of relative strengths (above average) as well as weaknesses (below average). The first 9 items are related to sustaining attention on tasks or play activities (e.g., "Engage in tasks that require sustained mental effort") while the other items assess hyperactivity and aggression. A recent study showed that the first nine items indeed captures one's ability to regulate attention (Saez, Folsom, Al Otaiba, & Schatschneider, 2012). Higher scores represent greater attentiveness. Teachers completed the SWAN checklist in the spring. Cronbach's alpha across the 9 items was .91.

Rapid Automatized Naming—The Letters subtest the Rapid Automatized Naming (Wolf & Denckla, 2005) was used. For this subtest, each examinee's completion time for naming a series of alternating lowercase letters was recorded. Test-retest reliability is .89 for children in elementary grades (Wolf & Dencckla, 2005)

Procedures

All assessments for the current study were conducted during the spring of the school year. Research assistants were trained prior to each assessment round, which consisted of two individual rounds and two small group sessions. Each research assistant spent approximately two hours in training and subsequent practice sessions for each round of assessments and was required to pass a fidelity check before administering assessments to the participants in order to ensure accuracy in administration and scoring. The trained research assistants assessed children individually during two sessions; the first session included the TOWRE, TNL Narrative Comprehension subtest, RAN, and WIAT Oral Reading Fluency and the second session included the WJ-III subtests and the OWLS. The order of assessments within each session varied across children in order to reduce fatigue effect. Then, all spelling and writing assessments were administered in small groups over two additional sessions. Throughout the assessments, students were given breaks as needed. Trained research assistants scored students' letter writing automaticity, story copying, spelling, and writing, and research assistants were trained to use each rubric on a small subset of the sample through practice and discussion of scoring issues.

Data Analysis Strategy

Primary analytic strategies were confirmatory factor analysis (CFA) and multilevel modeling. Latent variable approach (e.g., CFA) uses common variance among multiple indicators for a construct and thus reduces measurement error (Bollen, 1989; Kline, 2005). The first research question, dimensionality of writing, was examined using CFA. Assumptions (univariate and multivariate normality) were checked prior to analysis and were met. Model fits were evaluated using the following multiple indices: chi-square, comparative fit index (CFI), Tucker-Lewis index (TLI), root mean square error of approximation (RMSEA), and standardized root mean square residuals (SRMR). Differences in model fits for two nested models were evaluated by comparing chi-square differences between the two models. Confirmatory factor analysis was conducted using MPLUS 7 (Muthen & Muthen, 2012). Because children were nested within classroom and schools, the research questions 2 and 3 were addressed using 3 level multilevel modeling. PROC MIXED procedure of SAS 9.3 was used. Factor scores from CFA models (e.g., scores in the identified writing dimensions) were used in the multilevel modeling.

Results

Descriptive Statistics and Factor Analysis

Table 1 shows the means and standard deviations of writing scores by grade and gender. Where available, standard scores are presented. Note that the WIAT writing composition task was not normed for children in second grade, and thus, standard scores are not presented for this grade. In the WIAT writing composition, the standard score is a composite

of the standard scores from the number of words written and the theme development and organization standard scores. The standard score in the WIAT writing task is in the average range albeit slightly in the high average for children in grade 3 (mean standard score [SS] = 107.92, *SD* = 13.74). Standard scores in the WJ-III Writing Fluency task was in the average range as well (mean SS = 98.28 and 95.09 for grades 3 and 2, respectively). The WIAT Grammar Score, which is CIWS CBM writing, was in the average range (mean SS = 100.25) for students in grade 3. However, note that the standard scores for the Grammar Score should be taken with caution due to slight differences in scoring CIWS between WIAT and our approach following previous studies (e.g., McMaster et al., 2009).

Table 2 displays descriptive statistics for language and literacy predictors by grade and gender. In the language measures, mean performance was in the average range ranging from 8.65 on the TNL narrative comprehension to 99.13 on the WJ-III Picture Vocabulary task. Children's reading skills, spelling, and alphabet writing fluency were also in the average range. Correlations are presented in Table 3 for writing variables and in Table 4 for language and cognitive variables. Preliminary analysis showed that the patterns of relations were highly similar for children in grades 2 and 3, and thus, results from combined data are presented. The writing quality variables tended to be moderately and statistically significantly related to each other while writing productivity variables (number of words and number of ideas) were highly related to each other. Given that RAN has not been examined in relation to writing in previous studies with English-speaking children, correlations of RAN to writing scores are presented in Table 3. RAN was weakly to moderately related to all the writing variables ($-.24 \leq rs \leq -.43$). Language and cognitive variables in Table 4 were all statistically significantly correlated in expected directions.

Dimensionality of Writing

In order to examine the dimensionality captured in the various writing evaluation measures, we conducted a series of analysis. First, we confirmed the hypothesized factor structure using CFA models (measurement models) for the writing quality and productivity. Writing quality and productivity were deemed to be a good place to start because previous studies indicated that they are dissociable dimensions and their indicators are fairly well understood (Kim et al., 2014; Kim, Park, & Park, 2013; Puranik et al., 2008; Wagner et al., 2011). Second, we examined measurement model (i.e., CFA) of the CBM writing scores, and its relation to writing quality and writing productivity dimensions. Finally, we examined whether the WJ-III Writing Fluency is best described as an indicator of the writing quality, productivity, CBM writing, or as a separate observed variable.

We hypothesized that the theme and organization score of the WIAT composition task would capture the writing quality along with the idea development and organization aspects of the adapted 6+1 Trait rubric because the theme and organization of the WIAT task evaluates idea development and structural aspects of written composition. CFA confirmed the hypothesis: The model fit was good: χ^2 (13) = 72.92, *p* < .001; CFI = .95; TLI = .92; RMSEA = .097; and SRMR = .038. Factor loadings are presented in Table 1. Based on preliminary analysis, error covariance was allowed between theme and organization, and the 6+1 organization score. The CFA model for writing productivity using number of words

written and number of ideas yielded an excellent model fit: χ^2 (6) = 34.18, p < .001; CFI = .
99; TLI = .98; RMSEA = .10; and SRMR = .01.

To examine the dimensionality of the variables derived from the CBM scoring approaches,
two CFA models (two latent variables in which the CIWS variable is dissociable from
%CWS vs. one latent variable in which both CIWS and %CWS capture a single latent
variable) were fit and model fits were compared. The model fit for a single dimension was
slightly better (χ^2= 5.87, df = 1, p = .02). However, the CIWS and %CWS were very
highly correlated when modeled separately (r = .97). Therefore, it appeared reasonable to
model both the CWIS and %CWS as a single CBM latent variable (noted as CBM writing
scoring hereafter) in subsequent analysis. Table 5 shows comparison of CFA model fits for
alternative models examining whether writing quality, productivity, and CBM writing were
best considered as three dissociable variables or two dissociable variables, or as a single
variable. Results showed the three-latent-variable model describes the data best compared to
the other alternative models (χ^2 201.20, ps < .001).

Next, we examined whether the WJ-III Writing Fluency is best described as an indicator of
the identified dimensions of writing (writing quality, productivity, CBM) or is better
described as a separable variable. When we fit a model in which the WJ-III Writing Fluency
task was considered as a separate variable from the other three (i.e., writing quality,
productivity, and CBM writing), the fit was acceptable: χ^2 (151) = 1055.14, p < .001; CFI
= .90; TLI = .88; RMSEA = .011; SRMR = .08. The WJ-III Writing Fluency task correlated
most strongly with the writing quality at .67, followed by .59 with CBM writing, and .46
with productivity. When the WJ-III Writing Fluency task was considered as an indicator of
CBM writing or productivity, the model fits were statistically significantly worse (ps < .
001). When a CFA model was fit in which the WJ-III Writing Fluency was considered as an
indicator of writing quality, the model fit was not different from the separate dimension
model (i.e., the four factor model; χ^2 [df = 2] = 5.82, p = .054). Therefore, based on these
results and for parsimony, the WJ-III Writing Fluency task is considered as an indicator of
writing quality.

In summary, CFA analysis revealed the following three dimensions for the writing
outcomes: Writing quality, writing productivity, and CBM writing. The writing quality
dimension was strongly related to CBM writing at .82, and to writing productivity at .75.
Writing productivity and CBM writing were moderately correlated at .54.

Language and Cognitive Predictors of Writing Quality, Writing Productivity, and CBM Writing

Factor scores of the three writing dimensions (writing quality, productivity, and CBM
writing) from CFA results were extracted from MPLUS (SDs = 1.83, 25.74, and 28.74 for
writing quality, productivity, and CBM respectively; means are 0) and these three
dimensions were used in subsequent multilevel modeling using SAS 9.3. In addition, latent
variables were created for the predictors with multiple measures (i.e., oral language and
reading) using CFA. Factor loadings were high (see Table 2) and model fits were excellent
(not shown). Then, factor scores of these language and reading latent variables were used in
the multilevel models.

First, unconditional models without any predictors were fit for the three writing outcomes to parse out amount of variance attributable to individuals, classrooms, and schools. Intraclass correlations were as follows: (a) writing quality, .16 at the school level and .05 at the classroom level; (b) writing productivity, .07 at the school level, but 0 at the classroom level; (c) CBM writing, .16 at the school level and .15 at the classroom level. In other words, approximately 16% of the total variance in writing quality, 7% in writing productivity, and 16% in CBM writing were due to school differences whereas approximately 5% of the total variance in writing quality, 0% in writing productivity, and 15% were due to differences among classrooms. In the subsequent analysis, a three level model (school, classroom, and individual) was carried out for the writing quality and CBM outcomes whereas a two level model (school and individual) was conducted for the writing productivity outcome because of lack of variance at the classroom level in writing productivity.

We then fit models (M1) to examine unique correlates of writing quality, writing productivity, and CBM writing (research question 2). As shown in Tables 6 and 7, for writing quality, all the language and cognitive predictors were statistically significant after accounting for children's age: children's oral language (p = .004), reading (p < .001), spelling (p < .001), letter writing automaticity (p = .048), story copying (p < .001), RAN (p = .005), and attention (p = .03). After accounting for all these variables, no variance remained at the classroom and school levels. For the writing productivity, individual differences in reading (p = .002) and timed tasks such as letter writing automaticity (p = .004), story copying (p < .001), and RAN (p < .001) were related whereas oral language, spelling, and attention were not (ps ≥ .24). Finally, for the CBM writing scoring outcome, children's reading (p < .001), spelling (p < .001), story copying (p < .001), and attention (p = .02) remained statistically significant, whereas oral language, letter writing automaticity, and rapid automatized naming did not (ps ≥ .43).

Gender and Writing

To address the third research question of gender gap, first, children's gender was included as the main predictor in addition to the age control variable for each writing outcome. This allowed us to see whether gender differences were found after accounting for age, and if so, how large the gaps were before including any potential explanatory variables. As shown in the second models in Tables 6 and 7, in all the writing outcomes, boys had statistically significantly lower scores after accounting for age. In writing quality, boys had, on average, .39 standard deviations lower scores than girls. In writing productivity, boys' score was, on average, lower than girls by .46 standard deviations, and in CBM writing, boys' score was .37 standard deviations lower than girls.

Language and cognitive variables were then included in the models to investigate whether gender differences in writing score persisted or disappeared after controlling for these language and cognitive variables. Results in Tables 6 and 7 (M3) show that boys continued to have lower mean scores in writing even after accounting for all the included language and cognitive variables. However, the effect sizes were reduced by approximately quarter to a third compared to those in the initial models: the effect sizes were .22 in writing quality, .34 in writing productivity, and .22 in CBM writing. In other words, the included language and

literacy predictors explained the gender gap in writing outcomes to some extent, but the relation between gender and writing was not completely mediated by the included language and cognitive skills. It is of note that the relation of language and cognitive skills to the three writing outcomes essentially remained the same between M1 (before controlling for gender) and M3 (after accounting for gender). However, an exception was found for attention, which was no longer related to writing quality once gender was taken into consideration in addition to language and cognitive skills and age.

Discussion

In the present study, we investigated the dimensionality of writing, predictors of writing, and gender differences, using a large dataset from second and third grade students in the United States. Findings showed that writing quality, writing productivity, and CBM writing (CIWS and %CWS) were dissociable dimensions, at least for children in grades 2 and 3. Furthermore, unique predictors of each dimension differed.

In conjunction with previous studies (Kim et al., 2014; Puranik et al., 2008; Wagner et al., 2011), the present findings suggest that writing is not a single dimension, but is composed of multiple dimensions. Theoretically, the writing quality and productivity dimensions describe skills that are hypothesized to be products of two key components in writing, namely, ideation and transcription skills (Juel et al., 1986). Idea development and organization aspects, theme and organization scores in WIAT, and the WJ-III writing fluency task all captured the writing quality dimension whereas number of words written and number of ideas captured the writing productivity dimension. These findings confirm previous studies about the dissociability of writing quality and productivity (Kim et al., 2014; see also Puranik et al., 2008 and Wagner et al., 2011), but extend our understanding by demonstrating that the theme and organization score of WIAT and the sentence level WJ-III Writing Fluency tasks capture writing quality. It is interesting that the WJ-III Writing Fluency task was more strongly related to writing quality than to writing productivity or CBM writing, and was best described as an indicator of writing quality. This result suggests that the accuracy and rate at which children can construct sentences is likely to be an indicator of writing quality, but not writing productivity or CBM writing, at least at this stage of writing development. It might be that the WJ-III Writing Fluency task captures efficiency of children's transcription skills and sentence production skills (an oral language skill), both of which are important for written composition. It is plausible that this efficiency enables children to focus on higher order processes such as idea expression and organization.

Although CBM writing measures and coding methods have been examined for reliability and validity (see Graham et al., 2011; McMaster & Espin, 2007 for a review), the nature of their theoretical construct and dimensionality has been nebulous. In the present study, CBM writing scores captured a dissociable dimension from writing quality and productivity, and was strongly associated with the quality of writing at .82 and moderately associated with writing productivity at .54. It should be noted that in the present study, we included two scoring tools that are unique to CBM, CIWS and %CWS, although CBM writing scores also include other indicators such as total number of words written. This latter variable was

conceptualized as a productivity indicator in the present study according to previous studies and findings (e.g., Abbott & Berninger, 1993; Graham et al., 1997; Kim et al., 2014; Wagner et al., 2011). Whether the separate CBM writing dimension in the present study should be conceptualized as a global outcome measure of children's writing skill and/or writing fluency, or as another construct is beyond the scope of the present study. As noted earlier, CBM writing was recently theorized as writing fluency, which was defined as the ease to generate written text. According to automaticity and information processing theories (e.g., LaBerge & Samuels, 1974; Posner & Snyder, 1975), fluency (or automaticity) is required so that cognitive resources such as attention and working memory can be used for higher order cognitive resources. Applying this to writing development, efficiency in generating ideas and transcribing those ideas into written texts would allow a writer to focus on aspects such as presenting ideas in an organized, clear, and rich manner to enhance writing quality. The two CBM writing variables used in the present study (CIWS and %CWS) appear to operationalize writing fluency well because both capture not just the amount of writing, but efficiency (accuracy and amount). In addition, CIWS and %CWS tend to have highest validity evidence (e.g., Amato & Watkins, 2011; McMaster & Espin, 2007). One way to validate CBM writing measures (at least CIWS and %CWS) as indicators of writing fluency is to examine how data fit this theoretical hypothesis. Specifically, Ritchey et al. (in press) hypothesized that writing fluency includes text generation and transcription, which is aligned well with the simple view of writing (Juel et al., 1986) and the not-so-simple view of writing (Berninger & Winn, 2006). Therefore, text generation and transcription skills are component skills of writing fluency (i.e., CBM writing), which then would predict the criterion measure of writing such as writing quality. In other words, the CBM writing measures should mediate, at least partially, the relations of the text generation and transcription to the criterion measure of writing. Effort is under way to investigate this hypothesis by the current research team.

Another piece of evidence about multiplicity of writing dimensions comes from differential relations of language and cognitive skills to the three dimensions. In the model after accounting for gender (Models 3), whereas reading, letter writing fluency, and rapid automatized naming were related to both writing quality and productivity, oral language and spelling were related only to writing quality, but not to writing productivity. In addition, attention was related to the CBM writing outcome over and above the other variables in the model. Interestingly, although CIWS and %CWS do take into consideration grammatical accuracy, oral language skill did not uniquely influence the CBM writing. It is notable that reading was a consistent predictor for all three dimensions, underscoring the importance of early reading skill in early writing, even after accounting for other variables in the model. These results add to the increasing evidence of the relation between reading and writing, particularly in the elementary years (Berninger et al., 2002; Kim, 2013, 2014; Shanahan, 2006; Shanahan & Lomax, 1986). Reading has been hypothesized to play a role during the process of self-monitoring during planning and revision as children have to assess and plan for revisions (Hayes, 1996; McCuthen, Francis, & Kerr, 1997). Additionally, reading skills might contribute to the quality of writing by way of reading experiences – better readers read more and greater amount of reading might help children with idea generation with increased background knowledge and better organization of ideas (Berninger et al., 2006).

Transcription skills also tended to be consistently related to the writing outcomes. Spelling skill was related to writing quality and CBM writing, and letter writing automaticity was related to writing quality and writing productivity. These findings confirm previous studies about the role of transcription skills in writing, as they are needed not only to encode ideas into written language, but also to allow cognitive resources to be used for higher order writing processes (Abbott & Berninger, 1993; Berninger et al., 1997; Graham et al., 1997). It is noteworthy that compared to the letter writing task, the story copying task was related to all three of the writing outcomes after accounting for the other variables in the model, suggesting that story copying captures processes beyond those captured by the alphabet letter writing task. Story copying may involve a greater extent of processing capacity (e.g., working memory) to hold and process words and sentences as it is a discourse level text whereas a letter writing task is simply retrieval of letters from memory. Future studies are needed to replicate the results and any potential sources of differences between letter writing and story copying tasks.

Attention was another cognitive skill that was hypothesized to be important for writing (Berninger & Winn, 2006), and it was related to writing quality and CBM writing in the present study, confirming previous findings in first grade and in kindergarten (Kent et al., 2014; Kim et al., 2013). Interestingly, once children's gender was accounted for, attentiveness was not related to writing quality although its relation remained for the CBM writing outcome. These results suggest that gender may mediate the relation between attention and writing quality. Previous studies did not include gender as a covariate in examining the role of attention in writing. Future studies are needed to investigate the precise role of attention in writing development including reasons why attention matters for CBM writing. This is important for typically developing students, but also for students with ADHD, as boys are more commonly diagnosed than girls (Arcia & Conners, 1998; Levy, Hay, Bennett, & McStephen, 2005).

RAN was weakly to moderately related to various writing scores in bivariate correlations. Once other language and literacy skills were accounted for, RAN was independently related to writing quality and productivity, and to our knowledge, this was the first study to examine this relation in English. On the one hand, our findings converge with two previous studies in another orthography, Chinese (Chan et al., 2006; Ding et al., 2010). On the other hand, however, they are discrepant from a third study with Chinese children in which RAN was not related to writing once transcription skills were accounted for (Yan et al., 2012). If RAN captures mostly automaticity of letter retrieval, then its influence should be largely shared with handwriting automaticity tasks such as letter writing automaticity and story copying. Given its relations to writing quality and productivity, it appears that RAN captures processes beyond handwriting fluency. According to the multicomponent account of RAN (Wolf & Bowers, 1999; Wolf & Denkla, 2005), RAN includes processes for visual, orthographic, and verbal processing, and this integration process might be a factor that drives the independent relation of RAN to writing quality and productivity over and above the other language and literacy skills.

These results of multiple dimensions and associated predictors offer important implications for instruction and assessment practices. Instructionally, teachers may target different

aspects and skills to ensure student progress on all areas of writing. In addition, if data suggest that a child has weaknesses that may impact a particular writing dimension, teachers may target skills in that area of interest. For instance, if the teacher is mostly interested in improving children's writing quality, the teacher may want to model and introduce strategies to help students focus on the development and organization of ideas, and expressing generated ideas with appropriate language. Also it is worthy to note that to improve writing quality, instructional attention is needed in multiple aspects such as oral language, reading, transcription skills, RAN, and attention, given that the quality of writing was predicted by the wide array of language and literacy and cognitive assessments. If the teacher is particularly concerned about the child's productivity, the teacher may focus more on transcription related skills, given their roles in writing productivity. The teacher could target spelling, sentence writing fluency, or other related transcription skills.

Furthermore, if the teacher's primary goal is progress monitoring in writing, the CBM writing scores appear most appropriate for two reasons. First, although CBM and writing quality appear to be separable dimensions, CBM writing scores give a general idea about writing quality, given a strong relation between writing quality and CBM writing scores (r = .82). Second, CBM writing scores have been shown to be reliable and sensitive to growth captured within a relatively short span of time, which is important due to frequent assessments (e.g., two weeks; Graham et al., 2011; Lembke et al., 2003; McMaster & Espin, 2007; McMaster et al., 2009, 2011). In contrast, writing quality may be less appropriate for frequent assessments because writing quality indicators, which are typically evaluated on a rating scale, are not likely to be as sensitive as CBM writing measures in capturing changes during a short period. This speculation, however, requires a future study.

Finally, confirming previous studies (Berninger & Fuller; 1992; Knudson, 1995; NAEP), boys in the present study performed more poorly in all the three writing dimensions with effect sizes ranging from .37 to .46. Results further showed that gender differences were explained by the included language and cognitive skills as the effect sizes in gender differences were reduced by approximately quarter to a third when these variables were taken into account. In other words, the language and cognitive variables included in the present study partially explained writing performance differences between boys and girls. On the other hand, these results indicate that gender differences persisted in all of the three writing outcomes even after accounting for these language and cognitive skills. These findings indicate that studies are needed to expand our understanding of potential causes of gender gaps in writing. Given findings of a previous study that even in grade 1, boys engage in less writing (Graham et al., 2007), it would be informative to investigate how attitude toward writing together with language and literacy variables explains gender gaps in writing, and whether attitude is malleable. Additionally, other potential sources of gender gaps (e.g., persistence in writing; McKenna et al., 1995) need to be investigated in future studies.

Limitations and Conclusion

One of the limitations of the present study is that many children in the present study came from low-income family backgrounds from one mid-sized city in the Southeast In addition, the children were primarily African Americans and Whites, with virtually no English

language learners. Although their writing performance was in the average range in standardized and normed writing assessments, future research needs to determine whether similar results are found for children from different SES and linguistic backgrounds. Further understanding is also required regarding the CBM writing scoring dimension. Many studies have shown technical adequacy and the utility of CBM in screening and progress monitoring of elementary grade children's writing. Recent efforts in theoretical conceptualization (e.g., McMaster & Espin, 2007; Ritchey et al., in press) are in the right direction to help the field better understand this dimension of writing. Finally, there are other types of evaluative approaches to written compositions and predictors of writing skills which were not included in the present study. For instance, text elements (e.g., presence of text structural elements such as topic sentence, supporting details; Kulikowich et al., 2008; Wagner et al., 2011) and spelling and writing conventions (e.g., punctuation and handwriting) were not examined in the present study. In addition, motivational, discourse knowledge, and cognitive factors (e.g., strategic writing) have been shown to be related to writing skills (e.g., Bruning & Horn, 2000; Graham et al., 2005; Hidi & Boscolo, 2006; Limpo & Alves, 2013; Pajares, 2003; Olinghouse & Graham, 2009), but not examined in the present study.

Overall, the findings of the present study suggest that writing quality, writing productivity, and CBM writing (composed of CIWS and %CWS) are separate dimensions for children in grades 2 and 3, and that the relations of language and literacy variables differed for various writing outcomes. In addition, gender differences persist even after accounting for language and cognitive skills. Future research is needed to replicate the present study and to further expand our understanding about skills that influence children's writing development.

Acknowledgements

References

Abbott RD, Berninger VW. Structural equation modeling of relationships Among developmental skills and writing skills in primary- and intermediate-grade writers. Journal of Educational Psychology. 1993; 85:478–508.

Amato J, Watkins MW. The predictive validity of CBM writing indices of eighth-grade students. Journal of Special Education. 2011; 44:195–204.

Arcia E, Conners CK. Gender differences in ADHD? The Journal of Developmental and Behavioral Pediatrics. 1998; 19:77–83.

Bereiter, C.; Scardamalia, M. The Psychology of Written Composition. Lawrence Erlbaum Associates; Hillsdale, NJ: 1987.

Berman R, Verhoevan L. Cross-linguistic perspectives on the development of text-production abilities. Written Language and Literacy. 2002; 5:1–43.

Berninger VW. Coordinating transcription and text generation in working memory during composing: Automatized and constructive processes. Learning Disability Quarterly. 1999; 22:99–112.

Berninger VW, Abbott RD. Listening comprehension, oral expression, reading comprehension, and written expression: Related yet unique language systems in grades 1, 3, 5, and 7. Journal of Educational Psychology. 2010; 102:635–651. [PubMed: 21461140]

Berninger VW, Abbott RD, Abbott SP, Graham S, Richards T. Writing and reading: Connections between language by hand and language by eye. Journal of Learning Disabilities. 2002; 35:39–56. [PubMed: 15490899]

Berninger VW, Abbott RD, Jones J, Wolf BJ, Gould L, Anderson-Youngstrom M, Apel K. Early development of language by hand: Composing, reading, listening and speaking connections: Three-letter writing modes; and fast mapping in spelling. Developmental Neuropsychology. 2006; 29:61–92. [PubMed: 16390289]

Berninger VW, Abbott RD, Trivedi P, Olson E, Gould L, Westhaggen SY. Applying multiple dimensions of reading fluency to assessment and instruction. Journal of Psychoeducational Assessment. 2010; 28:3–18.

Berninger V, Fuller F. Gender differences in orthographic, verbal, and compositional fluency: implications for assessing writing disabilities in primary grade children. Journal of School Psychology. 1992; 30:363–382.

Berninger, VW.; Swanson, HL. Children's writing; toward a process theory of the development of skilled writing. In: Butterfield, E., editor. Children's writing: Toward a process theory of development of skilled writing. JAI Press; Greenwich, CT: 1994. p. 57-81.Stainthorp, R. Reproduced in The Learning and Teaching of Reading and Writing. Wiley: 2006.

Berninger VW, Vaughn KB, Graham S, Abbott RD, Abbott SP, Rogan LW, Brooks A, Reed E. Treatment of handwriting problems in beginning writers: Transfer from handwriting to composition. Journal of Educational Psychology. 1997; 89:652–666.

Berninger, VW.; Winn, WD. Implications of advancements in brain research and technology for writing development, writing instruction, and educational evolution. In: MacArthur, C.; Graham, S.; Fitzgerald, J., editors. Handbook of writing research. Guilford; New York: 2006. p. 96-114.

Bollen, KA. Structural Equations with Latent Variables. John Wiley & Sons, Inc.; New York: 1989.

Bowers PG. Tracing symbol naming speed's unique contributions to reading disabilities over time. Reading and Writing: An Interdisciplinary Journal. 1995; 7:189–216.

Bruning R, Horn C. Developing motivation to write. Educational Psychologist. 2000; 35:25–37.

Carrow-Woolfolk, E. Oral and written language scales. second edition. Western Psychological Services; Torrance, CA: 2011.

Casas AM, Ferrer MS, Fortea IB. Written composition performance of students with attention-deficit/hyperactivity disorder. Applied Psycholinguistics. 2013; 34:443–460.

Chan DW, Ho CSH, Tsang SM, Lee SH, Chung KKH. Exploring the reading-writing connection in Chinese children with dyslexia in Hong Kong. Reading and Writing: An Interdisciplinary Journal. 2006; 19:543–561.

Chenoweth NA, Hayes JR. The inner voice in writing. Written Communications. 2003; 20:99–118.

Coker DL, Ritchey KD. Curriculum based measurement of writing in kindergarten and first grade: An investigation of production and qualitative scores. Exceptional Children. 2010; 76:175–193.

Compton DL, DeFries JC, Olson RK. Are RAN and phonological awareness deficits additive in children with reading disabilities? Dyslexia. 2001; 7:125–149. [PubMed: 11765981]

Cragg L, Nation K. Exploring written narrative in children with poor reading comprehension. Educational Psychology. 2006; 26:55–72.

de Jong PF, van der Leij A. Developmental changes in the manifestation of a phonological deficit in dyslexic children learning to read a regular orthography. Journal of Educational Psychology. 2003; 95:22–40.

Deno SL. Curriculum-based measurement: The emerging alternative. Exceptional Children. 1985; 52:219–232. [PubMed: 2934262]

Ding Y, Richman LC, Yang L, Guo J. Rapid automatized naming and immediate memory functions in Chinese Mandarin-speaking elementary readers. Journal of Learning Disabilities. 2010; 43:48–61. [PubMed: 19770285]

Dockrell JE, Lindsay G, Connelly V. The impact of specific language impairment on adolescents' written text. Exceptional Children. 2009; 75:427–446.

Dockrell JE, Connelly V. The impact of oral language skills on the production of written text. BJEP Monograph Series ll, Number 6 – Teaching and Learning Writing. 2009; 1:45–62.

Dockrell JE, Connelly V. The role of oral language in underpinning the text generation difficulties in children with specific language impairment. Journal of Research in Reading. (in press).

Espin C, Shin J, Deno SL, Skare S, Robinson S, Benner B. Identifying indicators of written expression proficiency for middle school students. The Journal of Special Education. 2000; 34:140–153.

Espin CA, Weissenburger JW, Benson BJ. Assessing the writing performance of students in special education. Exceptionality. 2004; 12:55–66.

Gansle KA, VanDerHeyden AM, Noell GH, Resetar JL, Williams KL. The technical adequacy of curriculum-based and rating-based measures of written expression for elementary school students. School Psychology Review. 2006; 35:435–450.

Gillam, RB.; Pearson, NA. Test of Narrative Language. PRO-ED; Austin, TX: 2004.

Graham S. The role of production factors in learning disabled students' compositions. Journal of Educational Psychology. 1990; 82:781–791.

Graham S, Berninger VW, Abbott RD, Abbott SP, Whitaker D. Role of mechanics in composing of elementary school students: A new methodological approach. Journal of Educational Psychology. 1997; 89:170–182.

Graham S, Berninger VW, Fan W. The structural relationship between writing attitude and writing achievement in first and third grade students. Contemporary Educational Psychology. 2007; 32:516–536.

Graham S, Harris KR, Chorzempa BF. Contribution of spelling instruction to the spelling, writing, and reading of poor spellers. Journal of Educational Psychology. 2002; 94:669–686.

Graham S, Harris K, Fink B. Is handwriting casually related to learning to write? Treatment of handwriting problems in beginning writers. Journal of Educational Psychology. 2000; 92:620–633.

Graham, S.; Harris, KR.; Hebert, M. Informing writing: The benefits of formative assessment. Alliance for Excellent Education; Washington, DC: 2011.

Graham S, Harris KR, Mason L. Improving the writing performance, knowledge, self-efficacy of struggling young writers: The effects of self-regulated strategy development. Contemporary Educational Psychology. 2005; 30:207–241.

Gregg N, Coleman C, Stennett RB, Davis M. Discourse complexity of college writers with and without disabilities: A multidimensional analysis. Journal of Learning Disabilities. 2002; 35:23–38. [PubMed: 15490898]

Hayes, JR. Evidence from language bursts, revisions, and transcription for translation and its relation to other writing processes. In: Fayol, M.; Alamargot, D.; Berninger, V., editors. Translation of thought to written text while composing: Advancing theory, knowledge, methods, and applications. Psychology Press; East Sussex, UK: 2012. p. 45-67.

Hawke JL, Olson RK, Willcutt EG, Wadsworth SJ, DeFries JC. Gender ratios for reading difficulties. Dyslexia: An International Journal of Research and Practice. 2009; 15:239–242.

Hidi, S.; Boscolo, P. Motivation and writing. In: MacArthur, CA.; Graham, S.; Fitzgerald, J., editors. Handbooks of Writing Research. Guilford; New York, NY: 2006. p. 144-157.

Hooper SR, Costa L-J, McBee M, Anderson KL, Yerby DC, Knuth SB, Childress A. Concurrent and longitudinal neuropsychological contributors to written language expression in first and second grade students. Reading and Writing: An Interdisciplinary Journal. 2011; 24:221–252.

Hooper SR, Swartz CW, Wakely MB, de Kruif REL, Montomery JW. Executive functions in elementary school children with and without problems in written expression. Journal of Learning Disabilities. 2002; 35:57–68. [PubMed: 15490900]

Jones D, Christensen C. Relationship between automaticity in handwriting and students' ability to generate written text. Journal of Educational Psychology. 1999; 91:44–49.

Juel C, Griffith PL, Gough PB. Acquisition of literacy: A longitudinal study of children in first and second grade. Journal of Educational Psychology. 1986; 78:243–255.

Kail R, Hall LK. Processing speed, naming speed, and reading. Developmental Psychology. 1994; 30:949–954.

Kent S, Wanzek J, Petscher Y, Al Otaiba S, Kim Y-S. Writing fluency and quality in kindergarten and first grade: The role of attention, reading, transcription, and oral language. Reading and Writing: An Interdisciplinary Journal. 2014; 27(7):1163–1188.

Kim Y-S, Al Otaiba S, Folsom JS, Greulich L, Puranik C. Evaluating the dimensionality of first grade written composition. Journal of Speech, Language, and Hearing Research. 2014

Kim Y-S, Al Otaiba S, Folsom JS, Greulich L. Language, literacy, attentional behaviors, and instructional quality predictors of written composition for first graders. Early Childhood Research Quarterly. 2013; 28:461–469. [PubMed: 24062600]

Kim Y-S, Al Otaiba S, Puranik C, Folsom JS, Gruelich L, Wagner RK. Componental skills of beginning writing: An exploratory study at the end of kindergarten. Learning and Individual Differences. 2011; 21:517–525. [PubMed: 22267897]

Kim Y-S, Park C, Park Y. Is academic language use a separate dimension in begining writing? Evidence from Korean children. Learning and Individual Differences. 2013; 27:8–15.

Kim Y-S, Puranik C, Al Otaiba S. Developmental trajectories of writing skills in first grade: Examining the effects of SES and language and/or speech impairments. Elementary School Journal. 2014

Kirby J, Parrila RK, Pfeiffer SL. Naming speed and phonological awareness as predictors of reading development. Journal of Educational Psychology. 2003; 95:453–464.

Kline, RB. Principles and practice of structural equation modeling. 2nd ed. Guilford; New York, NY: 2005.

Knudson R. Development and application of a writing attitude survey for grades 1 to 3. Psychological Reports. 1992; 70:711–720. [PubMed: 1620761]

Knudson R. Writing experiences, attitudes, and achievement of first to sixth graders. Journal of Educational Research. 1995; 89:90–97.

Kulikowich JM, Mason LH, Brown SB. Evaluating fifth- and sixth-grade students' expository writing: task development, scoring, and psychometric issues. Reading and Writing: An Interdisciplinary Journal. 2008; 21:153–175.

LaBerge D, Samuels J. Towards a theory of automatic information processing in reading. Cognitive Psychology. 1974; 6:293–323.

Lee J. Can writing attitudes and learning behavior overcome gender difference in writing? Evidence from NAEP. Written Communication. 2013; 30:164–193.

Lembke E, Deno SL, Hall K. Identifying an indicator of growth in early writing proficiency for elementary school students. Assessment for Effective Intervention. 2003; 28:23–35.

Levy F, Hay DA, Bennett KS, McStephen M. Gender differences in ADHD subtype comorbidity. The Journal of the American Academy of Child and Adolescent Psychiatry. 2005; 44:368–76.

Limpo T, Alves RA. Modeling writing development: contribution of transcription and self-regulation to Portuguese students' text generation quality. Journal of Educational Psychology. 2013; 105:401–413.

Mackie C, Dockrell JE. The nature of written language deficits in children with SLI. Journal of Speech Language and Hearing Research. 2004; 47:1469–1483.

McGrew, KS.; Schrank, FA.; Woodcock, RW. Technical Manual. Woodcock-Johnson III Normative Update. Riverside Publishing; Rolling Meadows, IL: 2007.

McMaster K, Espin C. Technical features of curriculum-based measurement in writing: A literature review. The Journal of Special Education. 2007; 41:68–84.

McMaster KL, Du X, Pestursdottir AL. Technical features of curriculum-based measures for beginning writers. Journal of Learning Disabilities. 2009; 42:41–60. [PubMed: 19103799]

McMaster KL, Du X, Yeo S, Deno SL, Parker D, Ellis T. Curriculum-based measures of beginning writing: Technical features of the slope. Exceptional Children. 2011; 77:185–206.

Miles TR, Haslum MN, Wheeler TJ. Gender ratio in dyslexia. Annals of Dyslexia. 1990; 48:27–55.

National Center for Education Statistics. Persky, HR.; Dane, MC.; Jin, Y. The Nation's Report Card: Writing 2002, NCES 2003-529. 2003. Retrieved from http://nces.ed.gov/

National Center for Education Statistics. The Nation's Report Card: Reading 2011 (NCES 2012–457). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education; Washington, D.C.: 2011.

Posner, MI.; Snyder, CRR. Attention and cognition control. In: Solo, R., editor. Information processing and cognition: The Loyola symposium. Erlbaum Associates; Hillsdale, N. J.: 1975.

Olinghouse NG. Student- and instruction-level predictors of narrative writing in third-grade students. Reading and Writing: An Interdisciplinary Journal. 2008; 21:3–26.

Olinghouse NG, Graham S. The relationship between discourse knowledge and the writing performance of elementary-grade students. Journal of Educational Psychology. 2009; 101:37–50.

Pajares F. Self-efficacy beliefs, motivation, and achievement in writing: A review of the literature. Reading & Writing Quarterly. 2003; 19:139–158.

Pajares F, Valiante G. Grade level and gender differences in the writing self-beliefs of middle school students. Contemporary Educational Psychology. 1999; 24:390–405. [PubMed: 10508534]

Puranik CS, Lombardino LJ, Altmann LJ. Writing through retellings: An exploratory study of language-impaired and dyslexic populations. Reading and Writing: An Interdisciplinary Journal. 2007; 20:251–272.

Puranik, Cornoldi, Lombardino L, Altmann L. Assessing the microstructure of written language using a retelling paradigm. American Journal of Speech Language Pathology. 2008; 17:107–120. [PubMed: 18448599]

Re AM, Pedron M, Cornoldi C. Expressive writing. Difficulties in children described as exhibiting ADHD symptoms. Journal of Learning Disabilities. 2007; 40:244–255. [PubMed: 17518216]

Ritchey, KD.; McMaster, KL.; Al Otaiba, S.; Puranik, CS.; Kim, Y-S.; Parker, DC.; Ortiz, M. Indicators of fluent writing in beginning writers. In: Cummings, K.; Petscher, Y., editors. The Fluency Construct. Springer; New York City, NY: (in press)

Saez L, Folsom JS, Al Otaiba S, Schatschneider C. Relations among student attention behaviors, teacher practices, and beginning word reading skill. Journal of Learning Disabilities. 2012; 45:418–432. [PubMed: 22207616]

Savage RS, Frederickson N, Goodwin R, Patni U, Smith N, Tuersley L. Relationships among rapid digit naming, phonological processing, motor automaticity, and speech perception in poor, average, and good readers and spellers. Journal of Learning Disabilities. 2005; 38:12–28. [PubMed: 15727326]

Scardamalia, M.; Bereiter, C.; Goleman, H. The role of production factors in writing ability. In: Nystrand, M., editor. What writers know: The language, process, and structure of written discourse. Academic Press; San Diego, CA: 1982. p. 175-210.

Scott C, Windsor J. General language performance measures in spoken and written discourse produced by school-age children with and without language learning disabilities. Journal of Speech, Language, and Hearing Research. 2000; 43:324–339.

Shanahan, T. Relations among oral language, reading, and writing development. In: MacArthur, CA.; Graham, S., editors. Handbook of writing. Guilford Press; New York: 2006. p. 171-183.

Shanahan T, Lomax RG. An analysis and comparison of theoretical models of the reading–writing relationship. Journal of Educational Psychology. 1986; 78:116–123.

Shaywitz SE, Shaywitz BA, Fletcher JM, Escobar MD. Prevalence of reading disability in boys and girls. Journal of the American Medical Association. 1990; 264:998–1002. [PubMed: 2376893]

Shrout PE, Fleiss JL. Intraclass correlation: uses in assessing rater reliability. Psychological Bulletin. 1979; 86:420–428. [PubMed: 18839484]

Spring C, Davis JM. Relations of digit naming speed with three components of reading. Applied Psycholinguistics. 1988; 9:315–334.

Swanson, JM.; Schuck, S.; Mann, M.; Carlson, C.; Hartman, K.; Sergeant, JA.; McCleary, R. Categorical and dimensional definitions and evaluations of symptoms of ADHD: The SNAP and SWAN Rating Scales. 2006.

Torgesen, JK.; Wagner, RK.; Rashotte, CA. Test of word reading efficiency. second edition. PRO-ED; Austin, TX: 2012.

Wagner RK, Puranik CS, Foorman B, Foster E, Tschinkel E, Kantor PT. Modeling the development of written language. Reading and Writing: An Interdisciplinary Journal. 2011; 24:203–220.

Wagner RK, Torgesen JK. The nature of phonological processing and its causal role in the acquisition of reading skills. Psychological Bulletin. 1987; 101:192–212.

Wagner, RK.; Torgesen, JK.; Rashotte, CA.; Pearson, NA. Test of silent reading efficiency and comprehension. PRO-ED; Austin, TX: 2010.

Wechsler, D. Wechsler individual achievement test – third edition. Pearson; San Antonio, TX: 2009.

Wolf M, Bowers P. The double-deficit hypothesis for the developmental dyslexias. Journal of Educational Psychology. 1999; 91:415–438.

Wolf, M.; Denckla, MB. RAN/RAS: Rapid automatized naming and rapid alternating stimulus tests. PRO-ED; Austin, TX: 2005.

Wolf, M.; O'Brien, B. On issues of time, fluency and intervention. Fawcett, A., editor. Theory and good practice; Whurr; Dyslexia: London: 2001. p. 124-140.

Woodcock, RW.; McGrew, KS.; Mather, N. Woodcock-Johnson III tests of achievement. Riverside Publishing; Itasca, IL: 2001.

Yan CMW, McBride-Chang C, Wagner RK, Zhang J, Wong AMY, Shu H. Writing quality in Chinese children: speed and fluency matter. Reading and Writing: An Interdisciplinary Journal. 2012; 25:1499–1521.

Yoshmasu K, Barbaresi WJ, Colligan RC, Killian JM, Voigot RG, Weaver AL, Katusic SK. Gender, attention-deficit/hyperactivity disorder, and reading disability in a population-based birth cohort. Pediatrics. 2013; 131:637–644. [PubMed: 23460687]

Table 1

Means (Standard Deviations) of writing measures

	Grade 3			Grade 2			
	Entire Sample	Males	Females	Entire Sample	Males	Females	Loadings
WIAT total raw score (Words written + theme and organization)	88.82 (35.23)	81.90 (32.90)	97.55 (36.25)	79.89 (36.09)	71.35 (36.35)	87.21 (34.35)	NA
WIAT total score: SS	107.92 (13.74)	105.73 (14.06)	110.68 (12.87)	NA	NA	NA	NA
WIAT theme & organization raw	6.58 (2.84)	6.28 (2.90)	6.97 (2.74)	5.62 (2.59)	5.38 (2.65)	5.83 (2.53)	.72
WIAT theme & organization SS	105.09 (16.02)	103.71 (16.53)	106.83 (15.24)	NA	NA	NA	NA
WJ-III Writing Fluency raw	13.62 (4.41)	12.89 (4.11)	14.55 (4.62)	10.11 (4.97)	9.48 (4.92)	10.68 (4.96)	.67
WJ-III Writing Fluency SS	98.24 (15.78)	96.34 (13.85)	100.63 (17.69)	95.09 (26.54)	93.44 (24.29)	96.57 (28.44)	NA
Writing Quality Indicators							
WIAT Idea quality	3.89 (0.88)	3.81 (0.83)	3.98 (0.93)	3.44 (0.76)	3.32 (0.81)	3.55 (0.71)	.66
WIAT Organization	3.25 (0.89)	3.21 (0.86)	3.30 (0.94)	2.88 (0.82)	2.79 (0.86)	2.96 (0.79)	.70
Narrative Idea quality	4.46 (1.00)	4.30 (0.92)	4.66 (1.06)	4.10 (1.10)	3.99 (1.17)	4.19 (1.04)	.65
Narrative Organization	3.56 (0.87)	3.44 (0.76)	3.71 (0.97)	3.16 (0.78)	3.10 (0.84)	3.21 (0.73)	.63
Pet Idea quality	3.76 (0.80)	3.66 (0.76)	3.88 (0.83)	3.55 (0.81)	3.39 (0.80)	3.70 (0.80)	.54
Pet Organization	2.96 (0.69)	2.92 (0.71)	3.02 (0.67)	2.66 (0.69)	2.53 (0.65)	2.77 (0.69)	.60
CBM scores							
WIAT CWS	63.53 (31.85)	57.40 (30.30)	71.19 (32.27)	52.93 (29.97)	45.40 (28.15)	59.23 (20.11)	NA
WIAT IWS	26.43 (16.10)	25.47 (14.16)	27.71 (18.22)	29.55 (20.52)	27.94 (22.24)	30.91 (18.93)	NA
Narrative CWS	66.35 (35.35)	59.94 (31.00)	74.63 (38.89)	54.68 (28.01)	48.50 (25.65)	30.91 (18.93)	NA
Narrative IWS	32.36 (18.33)	31.63 (17.25)	33.30 (19.68)	37.19 (23.50)	34.82 (25.38)	39.16 (21.73)	NA
Pet CWS	62.87 (34.33)	54.31 (30.10)	73.22 (36.34)	54.10 (33.52)	45.25 (29.47)	61.69 (35.01)	NA
Pet IWS	25.57 (17.83)	24.91 (18.32)	26.36 (17.27)	27.40 (21.29)	26.25 (21.29)	28.39 (21.28)	NA
WIAT %CWS	.76 (.18)	.74 (.18)	.78 (.18)	.69 (.22)	.68 (.22)	.71 (.22)	.87
Narrative %CWS	.73 (.17)	.72 (.17)	.74 (.18)	.66 (.20)	.66 (.20)	.66 (.20)	.80

	Grade 3			Grade 2			
	Entire Sample	Males	Females	Entire Sample	Males	Females	Loadings
Pct %CWS	.78 (.20)	.75 (.21)	.81 (.19)	.72 (.22)	.69 (.22)	.74 (.22)	.78
WIAT CIWS	37.09 (35.02)	31.99 (33.18)	43.48 (36.35)	23.46 (34.10)	17.66 (32.66)	28.32 (34.65)	.87
WIAT CIWS SS	100.35 (17.13)	97.96 (16.55)	103.36 (17.44)	NA	NA	NA	NA
Narrative CIWS	33.99 (36.83)	28.31 (31.25)	41.33 (42.00)	17.48 (31.69)	13.67 (31.88)	20.64 (31.32)	.85
Pct CIWS	37.31 (36.71)	29.40 (31.85)	46.86 (39.94)	26.58 (34.84)	18.74 (29.86)	33.30 (37.43)	.79
Writing Productivity Indicators							
WIAT number of words	82.38 (33.66)	75.83 (31.17)	90.58 (34.97)	74.77 (34.99)	66.36 (34.90)	81.88 (33.59)	.87
WIAT number of words SS	109.00 (13.49)	106.51 (13.14)	112.15 (13.33)	NA	NA	NA	NA
Narrative number of words	89.44 (38.38)	82.01 (34.08)	99.03 (41.52)	82.36 (38.12)	73.79 (36.56)	89.47 (38.06)	.84
Pct number of words	80.05 (36.92)	72.52 (35.24)	89.14 (37.01)	73.86 (41.14)	64.33 (39.14)	82.11 (41.21)	.76
WIAT number of ideas	12.07 (5.04)	11.28 (4.80)	13.06 (5.18)	11.73 (5.51)	10.44 (5.34)	12.82 (5.43)	.80
Narrative number of ideas	15.62 (6.69)	14.30 (6.02)	17.31 (7.14)	14.26 (6.80)	12.76 (6.46)	15.50 (6.84)	.78
Pct number of ideas	12.95 (5.84)	11.74 (5.52)	14.41 (5.91)	11.69 (6.01)	10.47 (5.84)	12.76 (5.98)	.69

Note. WIAT = Wechsler Individual Achievement Test-Third Edition. SS = standard score. WJ-III = Woodcock Johnson III Tests of Achievement. CBM = curriculum-based measurement. CWS = correct word sequences. IWS = incorrect word sequences.

CIWS = correct minus incorrect word sequences.

Loadings were for the following latent variables: Writing quality, CBM writing, and productivity. The loadings of the WIAT theme and organization, and WJ-III Writing Fluency raw were those when they were considered as indicators of the writing quality.

Table 2

Means (Standard Deviations) of language and literacy predictors by gender

	Grade 3			Grade 2			Loadings
	Entire Sample	Males	Females	Entire Sample	Males	Females	
OWLS raw	87.09 (10.71)	87.29 (10.20)	86.84 (11.37)	82.54 (11.00)	81.15 (11.25)	83.74 (10.69)	.74
OWLS SS	98.09 (13.48)	98.12 (13.01)	98.05 (14.12)	101.40 (12.47)	99.57 (12.59)	102.98 (12.19)	NA
TNL Narrative Comprehension raw	28.34 (4.60)	27.91 (4.67)	28.87 (4.47)	26.60 (4.89)	25.79 (5.29)	27.29 (4.42)	.70
TNL Narrative Comprehension SS	8.65 (3.06)	8.33 (2.92)	9.04 (3.21)	8.33 (2.70)	7.86 (2.70)	8.73 (3.21)	NA
WJ-III Picture Vocabulary raw	23.24 (3.16)	23.22 (3.32)	23.28 (2.96)	21.50 (3.19)	21.58 (3.07)	21.43 (3.30)	.75
WJ-III Picture Vocabulary SS	99.13 (10.41)	99.00 (10.89)	99.30 (9.79)	98.74 (10.61)	98.87 (9.92)	98.64 (11.21)	NA
WJ-III LWID raw	50.26 (6.64)	50.04 (6.74)	50.54 (6.53)	44.37 (7.31)	43.84 (7.47)	44.82 (7.18)	.85
WJ-III LWID SS	104.84 (11.04)	104.40 (11.21)	105.41 (10.85)	105.89 (11.41)	104.84 (11.34)	106.80 (11.44)	NA
Sight Word Efficiency raw	62.88 (11.59)	61.79 (10.77)	64.20 (12.45)	54.45 (13.19)	53.43 (14.24)	55.31 (12.22)	.89
Sight Word Efficiency SS	96.26 (15.02)	94.50 (13.95)	98.43 (16.04)	99.57 (15.20)	98.05 (16.29)	100.84 (14.16)	NA
WIAT - ORF1	60.09 (23.48)	61.70 (23.14)	58.04 (23.86)	65.67 (28.27)	67.04 (29.86)	64.50 (26.89)	NA
WIAT - ORF2	71.69 (28.42)	74.06 (27.75)	68.70 (29.91)	77.83 (41.70)	83.45 (49.13)	73.02 (33.54)	NA
WIAT ORF Weighted raw	105.17 (35.81)	101.75 (34.41)	109.46 (37.21)	88.57 (33.43)	85.10 (33.81)	91.44 (32.96)	.88
WIAT ORF SS	103.49 (14.98)	101.85 (14.60)	105.54 (15.25)	99.23 (13.58)	97.37 (13.92)	100.78 (13.14)	NA
TOSREC raw	25.24 (9.27)	24.21 (9.39)	26.48 (9.02)	26.09 (9.75)	24.80 (9.90)	27.22 (9.52)	.79
TOSREC SS	101.22 (16.37)	99.36 (16.61)	103.48 (15.85)	98.64 (15.22)	96.64 (15.33)	100.37 (14.97)	NA
WJ-III Passage Comprehension raw	25.77 (3.49)	25.66 (3.52)	25.91 (3.47)	23.44 (3.92)	23.07 (3.95)	23.76 (3.87)	.80
WJ-III Passage Comprehension SS	95.38 (9.48)	95.05 (9.56)	95.82 (9.41)	97.51 (9.45)	96.44 (9.40)	98.45 (9.42)	NA
WIAT Alphabet Writing Fluency raw	17.57 (6.35)	17.32 (6.40)	17.87 (6.31)	15.91 (6.04)	15.76 (6.10)	16.03 (6.00)	NA
WIAT Alphabet Writing Fluency SS	104.74 (18.95)	104.12 (18.33)	105.53 (19.77)	104.62 (17.31)	104.22 (16.85)	104.97 (17.77)	NA
WJ-III Spelling raw	32.63 (5.88)	32.57 (5.88)	32.70 (5.90)	28.85 (5.51)	28.56 (5.88)	29.09 (5.20)	NA
WJ-III Spelling SS	102.75 (14.45)	102.50 (14.70)	103.06 (14.20)	102.58 (13.92)	101.46 (14.40)	103.53 (13.50)	NA

| | Grade 3 | | | | Grade 2 | | | |
	Entire Sample	Males	Females	Entire Sample	Males	Females	Loadings
Story copying: letters correct	36.21 (15.96)	33.51 (13.99)	39.60 (17.62)	27.55 (11.59)	26.17 (12.35)	28.75 (10.80)	NA
SWAN Attention	34.36 (10.13)	33.01 (9.73)	36.08 (10.42)	36.68 (11.90)	33.19 (11.77)	39.62 (11.24)	NA
RAN Time	28.40 (7.15)	28.45 (6.76)	28.35 (7.64)	32.53 (8.88)	33.12 (9.82)	32.03 (8.01)	NA
RAN SS	99.88 (12.63)	99.37 (11.98)	100.54 (13.43)	99.47 (12.84)	98.57 (13.24)	100.23 (12.49)	NA

Note. raw = raw score. OWLS = Oral and Written Language Scale. TNL = Test of Narrative Language. WJ-III = Woodcock Johnson-III Tests of Achievement. LWID = Letter Word Identification. SS = standard score. WIAT = Wechsler Individual Achievement Test-Third Edition. ORF = Oral Reading Fluency. TOSREC = Test of Silent Reading Efficiency and Comprehension.

RAN = rapid automatized naming

Loadings were for oral language latent variable (OWLS, TNL, & WJ picture vocabulary), and reading latent variable (WJ Letter Word Identification; TOWRE Sight Word Efficiency; Oral reading fluency; TOSREC; and WJ Passage Comprehension)

Table 3

Correlations among writing variables

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. WIAT TDTO	1																			
2. WIAT Q Ideas	.48	1																		
3. Narrative Q Ideas	.45	.48	1																	
4. Pet Q Ideas	.38	.44	.34	1																
5. WIAT Q Org	.57	.42	.40	.33	1															
6. Narrative Q Org	.43	.37	.66	.29	.43	1														
7. Pet Q Org	.38	.35	.40	.46	.43	.43	1													
8. WJ-III Writing Fluency	.46	.45	.42	.28	.48	.47	.39	1												
9. WIAT CIWS	.55	.52	.42	.40	.47	.43	.42	.53	1											
10. Narrative CIWS	.46	.42	.48	.37	.40	.49	.45	.48	.73	1										
11. Pet CIWS	.45	.42	.41	.53	.44	.42	.44	.44	.73	.69	1									
12. WIAT # of words	.47	.66	.40	.43	.26	.23	.23	.35	.50	.30	.39	1								
13. Narrative # of words	.44	.55	.58	.38	.22	.36	.27	.39	.42	.41	.40	.72	1							
14. Pet # of words	.36	.44	.31	.63	.20	.19	.27	.23	.33	.25	.50	.66	.63	1						
15. WIAT # of ideas	.37	.58	.36	.32	.17	.18	.15	.27	.38	.22	.28	.89	.68	.58	1					
16. Narrative # of ideas	.42	.54	.57	.34	.22	.37	.27	.38	.41	.39	.38	.68	.94	.57	.65	1				
17. Pet # of ideas	.35	.43	.35	.62	.23	.23	.32	.26	.34	.27	.50	.59	.58	.92	.54	.55	1			
18. WIAT %CWS	.44	.29	.33	.28	.41	.37	.42	.46	.85	.66	.61	.20	.22	.15	.14	.22	.19	1		
19. Narrative %CWS	.35	.19	.25	.21	.33	.32	.35	.35	.59	.83	.55	.08[ns]	.08[ns]	.06[ns]	.03[ns]	.09[ns]	.09[ns]	.71	1	
20. Pet %CWS	.35	.26	.35	.22	.43	.38	.38	.43	.62	.61	.78	.14	.20	.09	.07[ns]	.21	.14	.69	.65	1
21. RAN	−.36	−.37	−.40	−.29	−.30	−.32	−.29	−.43	−.40	−.36	−.37	−.41	−.43	−.34	−.31	−.42	−.33	−.35	−.24	−.35

Note. All the other coefficients are statistically significant at .05 level except those with [ns].

WIAT = Wechsler Individual Achievement Test-Third Edition. TDTO = Theme Development and Text Organization. Narrative = Narrative prompt; Pet = Pet prompt; CIWS = correct minus incorrect word sequences. CWS = correct word sequences.

Org = Organization. WJ-III = WJ-III = Woodcock Johnson III Tests of Achievement. Narrative = Narrative prompt. Q = Quality.

Table 4

Correlations among language and cognitive variables

	1	2	3	4	5	6	7	8	9	10	11	12
1. OWLS	1											
2. TNL Narrative Comprehension	.51	1										
3. WJ-III Picture Vocabulary	.55	.51	1									
4. WJ-III Letter Word Identification	.45	.40	.54	1								
5. TOWRE Sight Word Efficiency	.35	.34	.43	.76	1							
6. WIAT ORF Weighted	.40	.38	.47	.72	.80	1						
7. TOSREC	.42	.43	.48	.66	.69	.72	1					
8. WJ-III Passage Comprehension	.52	.46	.63	.74	.69	.67	.67	1				
9. WIAT Alphabet Writing Fluency	.21	.25	.24	.38	.39	.34	.35	.31	1			
10. WJ-III Spelling	.36	.28	.42	.78	.68	.66	.57	.59	.39	1		
11. Story copying : letters correct	.26	.25	.23	.35	.39	.40	.32	.36	.43	.41	1	
12. SWAN Attention	.36	.37	.32	.43	.44	.52	.59	.44	.26	.47	.47	1
13. RAN	-.16	-.17	-.18	-.48	-.69	-.52	-.41	-.43	-.35	-.46	-.27	-.36

Note. All coefficients are statistically significant at .001 level. OWLS: Oral and Written Language Scale. TNL = Test of Narrative Language. WJ-III = Woodcock Johnson-III Tests of Achievement. TOWRE = Test of Word Reading Efficiency. WIAT = Wechsler Individual Achievement Test-Third Edition. ORF = Oral Reading Fluency. TOSREC = Test of Silent Reading Efficiency and Comprehension. RAN = rapid automatized naming

Table 5

Model fit indices for alternative models

Model		χ^2 (df)	CFI	TLI	RMSEA	SRMR	Comparison to model 1: χ^2, df (p)
1	3 latent variables (Quality, Productivity, CBM)	1061.00 (153)	.90	.88	.11	.083	
2	2 latent variables (Quality+CBM, Productivity)	1262.21 (155)	.88	.85	.12	.098	201.21, 1 ($p < .001$)
3	2 latent variables (Productivity+CBM, Quality)	1702.65 (155)	.83	.79	.14	.12	641.65, 1 ($p < .001$)
4	2 latent variables (Quality +Productivity, CBM)	1344.67 (155)	.87	.84	.13	.106	283.67,1 ($p < .001$)
5	1 latent variable (Quality +Productivity+CBM)	1727.22 (156)	.83	.79	.14	.121	666.22, 2 ($p < .001$)

Note. CBM = curriculum-based measurement

Table 6

Results of multilevel models: Writing quality and writing productivity predicted by students' language and literacy skills, attention, and gender.

	Writing Quality			Writing Productivity		
Fixed effects	M1	M2	M3	M1	M2	M3
Intercept	−2.35 (.83)***	.34 (1.14)	−2.24 (.81)*	19.42 (16.28)	−3.11 (15.40)	21.71 (15.88)
Age in months	−.04 (.08)	−.02 (.14)	−.02 (.08)	−2.52 (1.59)	.96 (1.86)	−2.06 (1.55)
Male	NA	−.71 (.15)***	−.41 (.10)***	NA	−11.80 (2.22)***	−8.70 (1.93)***
Reading	.10 (.02)***		.10 (.02)***	.95 (.002)**		.91 (.30)**
Oral language	.03 (.01)**		.03 (.009)**	−.15 (.18)		−.11 (.17)
WJ-III spelling	.05 (.01)***		.06 (.01)***	.52 (.18)		−.13 (.24)
WIAT letter writing	.02 (.01)*		.02 (.009)*	.52 (.18)**		.55 (.17)**
Story copying	.03 (.004)***		.03 (.004)***	.54 (.08)***		.51 (.08)***
SWAN attention	.01 (.005)*		.008 (.006)	.13 (.11)		−.03 (.11)
RAN	−.02 (.008)**		−.02 (.008)**	−.73 (.15)***		−.75 (.15)***
Variance Components						
School	0	.57	0	17.42	43.16	15.65
Classroom	0	.25	0	NA	NA	NA
Children	1.07	2.42	1.03	366.44	580.09	349.77
−2LL	12.06.2	1929.9	1190.8	3641.8	4568.1	3621.8
AIC	1226.2	1941.7	1212.8	3663.8	4578.1	3645.8

Note: WJ-III = Woodcock Johnson-III Tests of Achievement; WIAT = Wechsler Individual Achievement Test; RAN = rapid automatized naming.

M1 examines the relation of language and cognitive skills; M2 examines the relation of gender and writing; and M3 examines the relation of gender and writing after accounting for language and cognitive skills.

Table 7

Results of multilevel models: CBM writing scoring predicted by students' language and literacy skills, attention, and gender.

	CBM writing		
Fixed effects	M1	M2	M3
Intercept	−68.45 (13.91)***	14.35 (17.89)	−66.39 (13.74)***
Age in months	−.67 (1.33)	−1.40 (2.55)	−.35 (1.32)
Male	NA	−10.67 (2.36)***	−6.35(1.70)***
Reading	1.48 (.27)***		1.45 (.26)***
Oral language	.12 (.16)		.16 (.15)
WJ-III spelling	−.04 (.16)		1.88 (.21)***
WIAT letter writing	1.81 (.22)***		−.02 (.15)
Story copying	.26 (.07)***		.24 (.07)***
SWAN attention	.30 (.09)***		.23 (.10)*
RAN	−.03 (.13)		−.04 (.13)
Variance Components			
School	5.12	150.82	6.09
Classroom	0	129.65	0
Children	284.10	528.60	274.05
−2LL	3528.5	4648.2	3514.8
AIC	3550.5	4660.2	3539.6

Note: WJ-III = Woodcock Johnson-III Tests of Achievement; WIAT = Wechsler Individual Achievement Test; RAN = rapid automatized naming.

M1 examines the relation of language and cognitive skills; M2 examines the relation of gender and writing; and M3 examines the relation of gender and writing after accounting for language and cognitive skills.